1. What is R studio?

R studio = IDE for R (simple and user-friendly)

- memory up - export plot - import dataset - manage script - create comanand etc.

3. Getting started with R.

① > x = 11 ② > y <- 7 ★くりかえしたら上書きされる ★消したら反応しない

> print(x) > y

> xx ← "~"

> yy ← "1"

a word not number

★ incomplete command は + と出る

★ 上下のキーを押すと前後のコマンドが表示される

★ #のあとは無視される

4. Create and Work with Vectors and Matrices in R

- vector → c (concatenate) を使う

ex. > x1 <- c(1,3,5,7,9)

> x1

> [1] 1 3 5 7 9

> rep(1, times = 10)

> [1] 1 1 1 1 1 1 1 1 1 1

> rep("marin", times = 5)

ex. > gender <- c("male", "female")

> 2:7

> seq(from = 1, to = 7, by = 1)

> rep(1:3, times = 5)

> rep(seq(from = 2, to = 5, by = 1), times = 5

> rep(c("m", "f"), times = 5)

- 2つのベクトルが同じ長さなら，対応した位置の数字等の要素を +-×÷ できる

- ベクトルの中の一つを抽出

ex. y = 1:5

y[3] とすると 3 が出る / y[-3] ⇒ 1,2,4,5 / y[1:3] ⇒ 1,2,3

y[c(1,5)] ⇒ 1,5 / y[-c(1,5)] ⇒ 2,3,4

y[y < 2.5] ⇒ 1,2

- matrix

> mat=matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=TRUE)

⇒ rowwise fashion

	[,1]	[,2]	[,3]
[1,]	1	2	3
[2,]	4	5	6
[3,]	7	8	9

⇒ FALSE にかえると

column-wise fashion

1	4	7
2	5	8
3	6	9

> mat ↗ これが出る

> mat[1,2] ⇒ 2 / mat[c(1,3),2] ⇒ 2,8 / mat[2,] ⇒ 4 5 6

> mat[,1] ⇒ 1 4 7 / mat*10

5. Import Data, Copy Data from Excel to R

☆ save the data file as ① .csv ② .txt better.

⇒ ダブルクリックで excel でひらくではない(?)

> help(read.csv)

> ? read.csv でる

(csv) > data1 ← read.csv(file.choose(), header = TRUE)

Tだけでよい

⇒ 出てきた window からファイルをえらぶ

> data1

> 表格

(csv) > data2 ← read.table(file.choose(), header = T, sep = ",")

> data2

>

(txt) > data3 ← read.delim(file.choose(), header = T)

> data4 ← read.table(file.choose(), header = T, sep = "\t")

5+ Using R studio to import excel data into R

① .xlsx ② .xls

file → import dataset → excel をえらぶ → browse → えらぶ → import

☆ missing value = 空白 or 1000 ⇒ NA のところで どう表示するか入力できる

code preview をコピー

→ 新しいファイルでひらく → コピペ → 実行 → reimport

6. Export Data from R

- write.table を使う

write.table(DataToExport, file = "ExportedFileName.csv", sep = ",")

↳ データ　↳ ファイル名

.txt

= "," 空白で区切る

= "\t" タブで区切る

row.names = F

current working directory にないとき

file = "/Users/ / / / ExportedFileName.csv",

name of path to new folder

- write.csv → sep = "," はいらない
- write.csv2

7. Import, check & work with data in R

```
> help (read.table)
```
データを読み込むとき

```
> ? read.table
```

Way①
```
> Data1 <- read.table(file = "データの位置", header = TRUE, sep = "\t")
```
Way②
```
> Data2 <- read.table(file.choose(), header = T, sep = "\t")
```
Way③
only for R studio Import dataset → from Text File → select → import
※quote → character を何で示すか

```
> dim(データ名) > [1] 725 6
```
↳ numbers of rows & columns

```
> head(データ名)
```
最初6列
```
> tail(データ名)
```
最後6列
) subset of data

```
> データ名[c(5,6,7,8,9), ]
[5:9, ]
[-(4:722), ]
```

```
> names(データ名)
> 性質の名前が現れる (Age, Height etc.)
```
(variable)

8. Import, check, work with data with R

× > mean(Age)

○ > mean(データ名$Age)

specify & extract data in "Age" variable (specific)) ① extract data

② attaching data
```
> attach(データ名)  ⟷  > detach(データ名)
> mean(Age)
> Age
```
) そのままできる (①ではできない)

```
> class(Variableの名前)
> "numeric" "integer" "factor"
```

```
> levels(variableの名前)
> "yes" "no" とか "female" "male"
"0" "1"
```

```
> summary(データ名)
```

```
> x <- c(0,1,1,1,0,0,0,0,0,0)  →  > x <- as.factor(x)
> class(x)                          > class(x)
> "numeric"                         > "factor"
> summary(x) → min median max など  > summary(x) → frequency : 0 1
                                                               7 3
```

9. Subsetting Data in R with Square Brackets

→ row / 行
column

↓
column / 列
row

> mean (Age [Gender == "female"])

> FemData <- データ名 [Gender == "female",]

> MaleOver15 <- データ名 [Gender == "male" & Age > 15,]

10. Logic Statement, cbind and rbind Functions

>Age [1:5]

> 6 18 16 14 5

> temp <- Age > 15

> temp [1:5

> F T T F F

> temp2 <- as.numeric (Age > 15)

>

[1] 0 1 1 0 0

> FemSmoke <- Gender == "female" & Smoke == "yes"

>

[1] F T F F F

> MoreData <- cbind (データ名, FemSmoke)

column-bind fashion

> rm (list = ls()) = Workspace 清除

11. Setting Up Working Directory in R

Session → SetWD → Choose Directory

> getwd()

> setwd ( )

> projectWD ( )

> setwd (projectWD)

12. Writing Scripts in R

tab → 可能なオーダーを表示してくれる

13. Installing Packages in R

> install.packages ("epiR") ()だけなら全部出せてくれる

> Selection = (A) 自分の地域番号

> remove.packages ( )

14. Customizing the Look of R studio

- Tools → Options

15. Apply Function in R

? apply

apply (X, MARGIN, FUN, ...)
- X: the object
- MARGIN: 1 = row, 2 = column どっちか
- FUN: function we want to apply
- ...: other arguments to send to function

ex. apply (X = StockData, MARGIN = 2, FUN = mean)
- MARGIN = 1 / = 2
- FUN: sum / plot, type = "l" / max
- , na.rm = TRUE : mean を計算するときに NA を除くということ

|| AVG だけでいい

colMeans (StockData, na.rm = TRUE) rowMeans ( ) もある

colMeans: column mean に特化

apply (X = StockData, MARGIN = 2, FUN = quantile, probs = c(0.2, 0.80), na.rm = TRUE)
- quantile: calculate percentile
- probs: which percentile to calculate

, FUN = plot, type = "l", main = "stock", ylab = "Price", xlab = "Day")

points (apply (X = StockData, MARGIN = 1, FUN = sum, na.rm = TRUE), pch = 16, col = "blue").

16. tApply Function in R

tApply: apply a function to subsets of variable or vector

tapply (X, INDEX, FUN = NULL, ..., simplify = TRUE)
- INDEX: grouping variable, same length with X, to create subsets of data

tapply (X = Age, INDEX = Smoke, FUN = Mean, na.rm = T) ○ だけでもよい

m = これの結果は
mean (Age[Smoke == "yes"])
"no" と同じ

, simplify = FALSE ⇒ list format rather than a simple vector

tapply (Age, Smoke, summary)
, quantile, probs = c(0.2, 0.8))

tapply || by

(X = Age, INDEX = list (Smoke, Gender), FUN = mean, na.rm = T)

↳ = mean (Age[Smoke == "no" & Gender == "female"])
"yes"   "male"